

Error-disturbance relations in mixed states

Masanao Ozawa

Graduate School of Information Science, Nagoya University, Chikusa-ku, Nagoya, 464-8601, Japan

Heisenberg's uncertainty principle was originally formulated in 1927 as a quantitative relation between the "mean error" of a measurement of one observable and the disturbance thereby caused on another observable. Heisenberg derived this famous relation under an additional assumption on quantum measurements that has been abandoned in the modern theory, and its universal validity was questioned in a debate on the sensitivity limit to gravitational-wave detectors in 1980s. A universally valid form of the error-disturbance relation was shown to be derived in the modern framework of general quantum measurements in 2003. We have experienced a considerable progress in theoretical and experimental study of error-disturbance relations in the last decade. In 2013 Branciard showed a new stronger form of universally valid error-disturbance relations, one of which is proved tight for spin measurements carried out in "pure" states. Nevertheless, a recent information-theoretical study of error-disturbance relations has suggested that Branciard relations can be considerably strengthened for measurements in mixed states. Here, we show a method for strengthening Branciard relations in mixed states and derive several new universally valid and stronger error-disturbance relations in mixed states. In particular, it is proved that one of them gives an ultimate error-disturbance relation for spin measurements, which is tight in any state. The new relations will play an important role in applications to state estimation problems including quantum cryptographic scenarios.

PACS numbers: 03.65.Ta, 06.20.Dk, 03.67.-a

Heisenberg's error-disturbance relation.—The discovery of quantum mechanics introduced non-commutativity in physical quantities: the commutation relation

$$[Q, P] = i\hbar \quad (1)$$

holds between a coordinate Q of a particle and its momentum P [72]. In 1927, Heisenberg found an operational meaning of the non-commutativity: "the more precisely the position is determined, the less precisely the momentum is known, and conversely [1, p. 64]." @ By the famous γ ray microscope thought experiment he derived the relation

$$\varepsilon(Q)\eta(P) \geq \frac{\hbar}{2}, \quad (2)$$

where $\varepsilon(Q)$ is the "mean error" [73] of a position measurement and $\eta(P)$ is the "discontinuous change" (disturbance) thereby caused on the momentum P [1, p. 64]. Heisenberg claimed that Eq. (2) is a "straightforward mathematical consequence" of Eq. (1) [1, p. 65] and gave its mathematical justification [1, p. 69], where Heisenberg derived the relation

$$\sigma(Q)\sigma(P) \geq \frac{\hbar}{2} \quad (3)$$

for the standard deviations $\sigma(Q), \sigma(P)$ [74] of the position Q and the momentum P [75]. Heisenberg applied Eq. (3) to the state just after the measurement assuming the approximate repeatability hypothesis: (AR) *Any measurement with error $\varepsilon(A)$ of an observable A leaves the object in the state satisfying $\sigma(A) \leq \varepsilon(A)$.* Under assumption (AR), Eq. (2) follows immediately from Eq. (3) [76].

Heisenberg's contemporaries supported assumption (AR) [77], as the "repeatability hypothesis" [78] was normally assumed for precise measurements. However, in the light of modern theory of quantum measurement [2–10], the "repeatability hypothesis" has been abandoned [79] and assumption

(AR) is no longer accepted. Thus, Eq. (2) cannot be considered as an immediate consequence of Eq. (3), although confusions have prevailed even in standard text books [11–14].

In 1980, Braginsky and coworkers [15] claimed that the Heisenberg error-disturbance relation (EDR) (2) leads to a sensitivity limit, called the *standard quantum limit* (SQL), for gravitational-wave detectors. Subsequently, Yuen [16] questioned the validity of the SQL and then Caves [17] defended the SQL by giving a new proof of the SQL without directly appealing to Eq. (2). Eventually, the conflict was reconciled by pointing out that Caves's derivation of the SQL still used (unfounded) assumption (AR), and a solvable model was constructed of an error-free position measurement that breaks the SQL [6, 18] (see also Ref. [19]). Later, this model was shown to also break the Heisenberg EDR (2) [20] with the uniquely determined notions of rms error and rms disturbance for quantum measurements [21]. Nowadays, the Heisenberg EDR (2) is taken to be a breakable limit [8, 22], but then the problem remains: what is the unbreakable constraint between error and disturbance, which Heisenberg originally intended?

Universally valid error-disturbance relations.—In 2003, the present author showed the relations

$$\varepsilon(A)\eta(B) + \varepsilon(A)\sigma(B) + \sigma(A)\eta(B) \geq \frac{1}{2}|\langle[A, B]\rangle|, \quad (4)$$

$$\varepsilon(A)\eta(B) + |\langle[n(A), B]\rangle + \langle[A, d(B)]\rangle| \geq \frac{1}{2}|\langle[A, B]\rangle|, \quad (5)$$

which are universally valid for any observables A, B , any system state, and any measuring apparatus, where standard deviations $\sigma(\cdots)$ and expectation values $\langle\cdots\rangle$ are taken in the state just before measurement, and $n(A)$ and $d(B)$ are system observables representing the first moments of the error and the disturbance for A and B , respectively [10, 23–27]. Relation (5) concludes that if the error and the disturbance are statistically independent from system state, then the Heisen-

berg EDR

$$\varepsilon(A)\eta(B) \geq \frac{1}{2} | \langle [A, B] \rangle | \quad (6)$$

holds, extending the previous results [28–31]. Relation (4) leads to the following new constraints for error-free measurements and non-disturbing measurements: if $\varepsilon(A) = 0$ then

$$\sigma(A)\eta(B) \geq \frac{1}{2} | \langle [A, B] \rangle |, \quad (7)$$

and if $\eta(B) = 0$ then

$$\varepsilon(A)\sigma(B) \geq \frac{1}{2} | \langle [A, B] \rangle |, \quad (8)$$

in contrast to that the Heisenberg EDR (6) leads to the divergence of $\varepsilon(A)$ or $\eta(A)$ if $\langle [A, B] \rangle \neq 0$. Relation (8) has been used to derive conservation-law-induced limits for measurements [25, 32] (see also [33, 34]), quantitatively generalizing the Wigner-Araki-Yanase no-go theorem [35–38] for repeatable measurements under conservation laws. Moreover, it has been used to derive an accuracy limit for quantum computing induced by conservation laws [25] (see also [39–45]).

To derive the above relations, the “mean error” $\varepsilon(A)$ and the disturbance $\eta(B)$ are defined as follows. Let us consider a (model of) measuring process $(\mathcal{K}, |\xi\rangle, U, M)$ for a system \mathbf{S} described by a Hilbert space \mathcal{H} determined by the probe system \mathbf{P} described by a Hilbert space \mathcal{K} , the initial probe state $|\xi\rangle$, the unitary evolution U of the composite system $\mathbf{S} + \mathbf{P}$ during the measuring interaction, and the meter observable M of the probe \mathbf{P} to be directly observed [4]. Then, for observables A, B of \mathbf{S} , the *error observable* $N(A)$ and the *disturbance observable* $D(B)$ are defined by

$$N(A) = M(\Delta t) - A(0), \quad (9)$$

$$D(B) = B(\Delta t) - B(0), \quad (10)$$

where $A(0) = A \otimes I_{\mathcal{K}}$, $B(0) = B \otimes I_{\mathcal{K}}$, $M(0) = I_{\mathcal{H}} \otimes M$, $B(\Delta t) = U^* B(0) U$, and $M(\Delta t) = U^* M(0) U$. Then, the (root-mean-square) error $\varepsilon(A)$ and the (root-mean-square) disturbance $\eta(B)$ are defined by [24]

$$\varepsilon(A)^2 = \text{Tr}[N(A)^2 \rho \otimes |\xi\rangle\langle\xi|], \quad (11)$$

$$\eta(B)^2 = \text{Tr}[D(B)^2 \rho \otimes |\xi\rangle\langle\xi|]. \quad (12)$$

Branciard’s relations.—In 2013, Branciard [46] discussed the tightness of Eq. (4) and improved Eq. (4) as

$$\varepsilon(A)^2 \sigma(B)^2 + \sigma(A)^2 \eta(B)^2 + 2\varepsilon(A)\eta(B) \sqrt{\sigma(A)^2 \sigma(B)^2 - C_{AB}^2} \geq C_{AB}^2, \quad (13)$$

which is universally valid and stronger than Eq. (4) in the case where $\varepsilon(A) \neq 0$ or $\eta(B) \neq 0$, where

$$C_{AB} = \frac{1}{2i} \text{Tr}([A, B] \rho). \quad (14)$$

From Eq. (13) he showed that the equality in Eq. (4) cannot be attained unless $\varepsilon(A)\eta(B) = 0$. It was also shown [46] that the above relation can be further strengthened to be a tight relation for spin measurements in pure states as follows. Let A and B be 2-valued observables with eigenvalues ± 1 and let ρ be a state possibly mixed for which $\langle A \rangle = \langle B \rangle = 0$, and we further suppose the *same spectrum condition* that the meter M has the same spectrum as the measured observable A . In this case, Branciard [46] derived the relation

$$\hat{\varepsilon}(A)^2 + \hat{\eta}(B)^2 + 2\hat{\varepsilon}(A)\hat{\eta}(B) \sqrt{1 - C_{AB}^2} \geq C_{AB}^2, \quad (15)$$

where

$$\hat{\varepsilon}(A) = \varepsilon(A) \sqrt{1 - \frac{\varepsilon(A)^2}{4}}, \quad \hat{\eta}(B) = \eta(B) \sqrt{1 - \frac{\eta(B)^2}{4}}, \quad (16)$$

more stringent than Eq. (13) and showed its tightness for pure input states.

Problem of mixed states.—In recent papers [47, 48], it is suggested that relation (13) would be considerably strengthened if the measured system is in a mixed state. For example, consider the case of spin measurement where $A = Z$, $B = X$ for Pauli operators X, Y, Z of a spin 1/2 system \mathbf{S} , the meter M has the same spectrum as A , and the input state is completely random, i.e., $\rho = I_{\mathcal{H}}/2$. In this case, Eq. (13) gives no constraint, since $C_{AB} = 0$, but a result from an information theoretical approach [47] leads to the relation

$$\left[\varepsilon(Z)^2 + \frac{1}{3} \right] \left[\eta(X)^2 + \frac{1}{3} \right] \geq \frac{16}{\pi^2 e^2} \approx 0.219. \quad (17)$$

In particular, further consideration concludes the relation $\eta(B) = \sqrt{2}$ if $\varepsilon(A) = 0$.

In what follows, we show a method for strengthening Branciard relations in mixed states and derive several new universally valid and stronger error-disturbance relations in mixed states. In particular, it is proved that one of them gives an ultimate error-disturbance relation for spin measurements, which is tight in any state.

Purification.—Suppose that the input state ρ has eigenvalues $p_j > 0$ with spectral decomposition $\rho = \sum_j p_j |\phi_j\rangle\langle\phi_j|$. Then, for any Hilbert space \mathcal{H}' with $\dim(\mathcal{H}') \geq \text{rank}(\rho)$ describing an ancillary system, we have a “purification” $|\Psi\rangle \in \mathcal{H} \otimes \mathcal{H}'$ of ρ with the Schmidt decomposition $|\Psi\rangle = \sum_j \sqrt{p_j} |\phi_j\rangle \otimes |\eta_j\rangle$, which satisfies $\rho = \text{Tr}_{\mathcal{H}'}[|\Psi\rangle\langle\Psi|]$, where $\{|\eta_j\rangle\}$ is an arbitrary orthonormal family in \mathcal{H}' and $\text{Tr}_{\mathcal{H}'}$ stands for the partial trace over \mathcal{H}' . In this case, we can extend A, B to $\mathcal{H} \otimes \mathcal{H}'$ by $A' = A \otimes I_{\mathcal{H}'}$ and $B' = B \otimes I_{\mathcal{H}'}$, and extend any measuring process $(\mathcal{K}, |\xi\rangle, U, M)$ for \mathcal{H} to a measuring process $(\mathcal{K}, |\xi\rangle, U', M)$ for $\mathcal{H} \otimes \mathcal{H}'$ by $U' = U \otimes I_{\mathcal{H}'}$. Then, we easily obtain

$$\sigma(A') = \sigma(A), \quad \sigma(B') = \sigma(B), \quad (18)$$

$$\varepsilon(A') = \varepsilon(A), \quad \eta(B') = \eta(B), \quad C_{A'B'} = C_{AB}. \quad (19)$$

Hence, the above extensions preserve relation (13). Then, it is easily understood that in the mixed state case Eq. (13) is

less stringent than the pure state case, since the extended measuring apparatus $(\mathcal{K}, |\xi\rangle, U', M)$ does not interact with the ancillary system described by \mathcal{H}' in the measured system [48].

Strengthening of error-disturbance relations for mixed states.—In the above discussion, the extension conserves all quantities $\sigma(A)$, $\sigma(B)$, $\varepsilon(A)$, $\eta(B)$, and C_{AB} . Now, we consider another extension using the “canonical purification” that leads to a stronger relation for the quantities $\sigma(A)$, $\sigma(B)$, $\varepsilon(A)$, and $\eta(B)$.

For this purpose a particularly important choice of the ancillary Hilbert space \mathcal{H}' is the dual space \mathcal{H}^* of \mathcal{H} , which consists of all bra vectors $\langle\psi| \in \mathcal{H}^*$. Now we suppose $\mathcal{H}' = \mathcal{H}^*$. Then, we define the “canonical purification” $|\Psi\rangle \in \mathcal{H} \otimes \mathcal{H}^*$ of ρ by

$$|\Psi\rangle = \sum_j \sqrt{p_j} |\phi_j\rangle \otimes \langle\phi_j|, \quad (20)$$

which satisfies $\rho = \text{Tr}_{\mathcal{H}^*}[|\Psi\rangle\langle\Psi|]$. Let

$$-i\sqrt{\rho}[A, B]\sqrt{\rho} = W^\dagger |\sqrt{\rho}[A, B]\sqrt{\rho}| \quad (21)$$

be a polar decomposition of the operator $-i\sqrt{\rho}[A, B]\sqrt{\rho}$. Since $-i\sqrt{\rho}[A, B]\sqrt{\rho}$ is self-adjoint, W can be chosen as a self-adjoint unitary operator on \mathcal{H} . Then, the corresponding operator W^* on \mathcal{H}^* also is a self-adjoint unitary operator. We extend the observable B on \mathcal{H} to the observable B'_W on $\mathcal{H} \otimes \mathcal{H}^*$ by

$$B'_W = (B - \text{Tr}[B\rho]I_{\mathcal{H}}) \otimes W^*. \quad (22)$$

Now, we define a new constant D_{AB} by

$$D_{AB} = \frac{1}{2} \text{Tr}(|\sqrt{\rho}[A, B]\sqrt{\rho}|). \quad (23)$$

Then, as shown in Supplemental Material we have

$$\begin{aligned} \sigma(A') &= \sigma(A), \quad \sigma(B'_W) \leq \sigma(B), \\ \varepsilon(A') &= \varepsilon(A), \quad \eta(B'_W) = \eta(B), \quad C_{A'B'_W} = D_{AB}. \end{aligned} \quad (24)$$

Therefore, from Eq. (13) we obtain

$$\begin{aligned} &\varepsilon(A)^2 \sigma(B)^2 + \sigma(A)^2 \eta(B)^2 \\ &+ 2\varepsilon(A)\eta(B)\sqrt{\sigma(A)^2 \sigma(B)^2 - D_{AB}^2} \geq D_{AB}^2. \end{aligned} \quad (26)$$

This relation is equivalent to Eq. (13) if ρ is a pure state, but significantly stronger than Eq. (13) if ρ is a mixed state; for $\rho = I_{\mathcal{H}}/2$, $A = Z$, and $B = X$, we have $C_{AB} = 0$ but $D_{AB} = 1$.

In other words, we consider the physically same measurement $(\mathcal{K}, |\xi\rangle, U', M)$ as the original model $(\mathcal{K}, |\xi\rangle, U, M)$, but we consider the disturbance on a new observable B'_W physically different from the original observable $B' = B \otimes I_{\mathcal{H}'}$. Then, B'_W has the mathematically same value of disturbance $\eta(B'_W)$ as the original value $\eta(B) = \eta(B')$, and $\sigma(B'_W)$ is favorably smaller than the original value $\sigma(B) = \sigma(B')$. Nevertheless, the lower bound $C_{A'B'}$ can be larger than the original value C_{AB} , which can be maximized up to D_{AB} to obtain

the significantly stringent relation, Eq. (26), for the original values $\sigma(A)$, $\sigma(B)$, $\varepsilon(A)$, and $\eta(B)$.

Error-disturbance relation for binary measurements.—As the simplest choice of A and B , the error and disturbance in spin measurements have been extensively studied in theoretically and experimentally [49–55]. Let us consider 2-valued observables A, B with eigenvalues ± 1 and a state ρ possibly mixed for which $\langle A \rangle = \langle B \rangle = 0$, and we further suppose the *same spectrum condition* that the meter M has the same spectrum as the measured observable A . In this case, Branciard [46] derived Eq. (15), which is more stringent than Eq. (13) and showed its tightness for pure input states.

In this particular case (i.e. $A^2 = B^2 = I_{\mathcal{H}}$, $M^2 = I_{\mathcal{K}}$, and $\text{Tr}[A\rho] = \text{Tr}[B\rho] = 0$), as shown in Supplemental Material the canonical purification enables us to strengthen Branciard relation (15) to the relation

$$\hat{\varepsilon}(A)^2 + \hat{\eta}(B)^2 + 2\hat{\varepsilon}(A)\hat{\eta}(B)\sqrt{1 - D_{AB}^2} \geq D_{AB}^2, \quad (27)$$

which is equivalent to Eq. (15) if ρ is a pure state, but considerably stronger than Eq. (15) if ρ is a mixed state. This strengthening for mixed states will play an important role in applications to state estimation problems including quantum cryptographic scenarios.

Tight error-disturbance relation for spin measurements.—Now we shall show the tightness of Eq. (27) for any state ρ in spin measurements. Let $\mathcal{H} \cong \mathbb{C}^2$ and let X, Y , and Z be Pauli operators thereon. We denote by $|0\rangle$ and $|1\rangle$ the eigenstates of Z with eigenvalues $+1$ and -1 , respectively. Let $A = Z$ and $B = X$. Let us consider a measuring process $(\mathcal{K}, |\xi\rangle, U, M)$ for \mathcal{H} satisfying the same spectrum condition $M^2 = I_{\mathcal{K}}$. Suppose that the input state ρ satisfies $\text{Tr}[A\rho] = \text{Tr}[B\rho] = 0$. In this case, we have $\rho = \frac{1}{2}(\alpha Y + I_{\mathcal{H}})$ with $-1 \leq \alpha \leq 1$. Hence, we have $D_{AB} = 1$.

Thus, from Eq. (27) the relation

$$(\varepsilon(Z)^2 - 2)^2 + (\eta(X)^2 - 2)^2 \leq 4 \quad (28)$$

holds for every measuring apparatus satisfying the same spectrum condition.

For the case $\rho = I_{\mathcal{H}}/2$, we have $C_{AB} = 0$ and the Branciard relation, Eq. (15), leads to no constraint. It can be easily seen that Eq. (28) is significantly stronger than previous relation Eq. (17) obtained from information theoretical approach [47].

In particular, we have the relation $\eta(X) = \sqrt{2}$, if $\varepsilon(Z) = 0$, which was also previously obtained in [47], so that the precise measurement is attained only with the maximum disturbance, and conversely we have $\varepsilon(Z) = \sqrt{2}$ if $\eta(X) = 0$, so that we conclude *no information from no disturbance*.

Now we shall show that Eq. (28) is attained by a measuring process uniformly for any state ρ . Suppose that the probe is another spin 1/2 system described by the Hilbert space $\mathcal{K} \cong \mathbb{C}^2$. We take the initial probe state as $|\xi\rangle = |0'\rangle$ and the meter observable in the probe as $M = Z'$; for distinction the prime indicates what defined for \mathcal{K} . The unitary operator U for the

measuring interaction is given by

$$U = C[X'](I \otimes W(\theta)), \quad (29)$$

where controlled not $C[X']$ and rotation $W(\theta)$ are given by

$$C[X'] = |0\rangle\langle 0| \otimes I_K + |1\rangle\langle 1| \otimes X', \quad (30)$$

$$W(\theta) = \cos \theta Z' + \sin \theta X'. \quad (31)$$

Then, the rms error $\varepsilon(Z)$ and the rms disturbance $\eta(X)$ of the measuring process $(C^2, |0'\rangle, U, Z')$ are obtained as

$$\varepsilon(Z)^2 = 4 \sin^2 \theta, \quad \eta(X)^2 = 4 \sin^2 \left(\frac{\pi}{4} - \theta \right) \quad (32)$$

for every input state ρ [52]. It follows that we have

$$(\varepsilon(Z)^2 - 2)^2 + (\eta(X)^2 - 2)^2 = 4. \quad (33)$$

Thus, Eq. (28) is attained by $(C^2, |0'\rangle, U, Z')$ with U defined by Eq. (29) uniformly for any input state ρ .

Most general error-tradeoff relation.—Now we shall consider the most general form of the error-disturbance relation. For this purpose it is convenient to introduce more general and simpler type of mathematical models. A *joint measurement model* for a Hilbert space \mathcal{H} is determined by a quadruple $(\mathcal{K}, |\xi\rangle, \mathcal{A}, \mathcal{B})$ consisting of a Hilbert space \mathcal{K} , a unit vector $|\xi\rangle \in \mathcal{K}$, and mutually commuting self-adjoint operators \mathcal{A}, \mathcal{B} on $\mathcal{H} \otimes \mathcal{K}$. For any pair of observables A, B and a density operator ρ on \mathcal{H} , the rms errors $\varepsilon(A)$ and $\varepsilon(B)$ for joint A, B measurement by $(\mathcal{K}, |\xi\rangle, \mathcal{A}, \mathcal{B})$ in ρ are defined by

$$\varepsilon(A, \mathcal{A}, \rho) = \text{Tr}[(\mathcal{A} - A \otimes I_K)^2 \rho \otimes |\xi\rangle\langle\xi|]^{1/2}, \quad (34)$$

$$\varepsilon(B, \mathcal{B}, \rho) = \text{Tr}[(\mathcal{B} - B \otimes I_K)^2 \rho \otimes |\xi\rangle\langle\xi|]^{1/2}. \quad (35)$$

Every measuring process $(\mathcal{K}, |\xi\rangle, U, M)$ with $\varepsilon(A)$ and $\eta(B)$ is a joint measurement model $(\mathcal{K}, |\xi\rangle, \mathcal{A}, \mathcal{B})$ with $\mathcal{A} = M(\Delta t)$, $\mathcal{B} = B(\Delta t)$, $\varepsilon(A, \mathcal{A}, \rho) = \varepsilon(A)$, and $\varepsilon(B, \mathcal{B}, \rho) = \eta(B)$. From now on, we abbreviate $\varepsilon(A) = \varepsilon(A, \mathcal{A}, \rho)$ and $\varepsilon(B) = \varepsilon(B, \mathcal{B}, \rho)$ when confusions may not occur.

The joint measurement model $(\mathcal{K}, |\xi\rangle, \mathcal{A}, \mathcal{B})$ also defines the standard deviations $\sigma(\mathcal{A}), \sigma(\mathcal{B})$ in the state $\rho \otimes |\xi\rangle\langle\xi|$ and the (first moment) biases $\delta(A) = \text{Tr}[(\mathcal{A} - A \otimes I_K)\rho \otimes |\xi\rangle\langle\xi|]$, $\delta(B) = \text{Tr}[(\mathcal{B} - B \otimes I_K)\rho \otimes |\xi\rangle\langle\xi|]$. Assuming $\sigma(\mathcal{A}), \sigma(\mathcal{B}) \neq 0$, define

$$\begin{aligned} E_{\sigma(\mathcal{A}), \delta(A)}(A) &= \sqrt{\sigma(A)^2 - \left(\frac{\sigma(A)^2 + \sigma(\mathcal{A})^2 - (\varepsilon(A)^2 - \delta(A)^2)}{2 \sigma(\mathcal{A})} \right)^2}, \\ E_{\sigma(\mathcal{B}), \delta(B)}(B) &= \sqrt{\sigma(B)^2 - \left(\frac{\sigma(B)^2 + \sigma(\mathcal{B})^2 - (\varepsilon(B)^2 - \delta(B)^2)}{2 \sigma(\mathcal{B})} \right)^2}. \end{aligned}$$

Then, as shown in Supplemental Material we obtain the relation

$$\begin{aligned} &E_{\sigma(\mathcal{A}), \delta(A)}(A)^2 \sigma(B)^2 + \sigma(A)^2 E_{\sigma(\mathcal{B}), \delta(B)}(B)^2 \\ &+ 2 E_{\sigma(\mathcal{A}), \delta(A)}(A) E_{\sigma(\mathcal{B}), \delta(B)}(B) \sqrt{\sigma(A)^2 \sigma(B)^2 - D_{AB}^2} \\ &\geq D_{AB}^2. \end{aligned} \quad (36)$$

for any given values of $\sigma(\mathcal{A}) > 0, \sigma(\mathcal{B}) > 0$, and $\delta(A), \delta(B)$, which for mixed states strengthens the corresponding relation with C_{AB} recently obtained by Branciard [48]. Branciard [48] showed that his relation is stronger than the universally valid error-tradeoff relations previously obtained by the present author [24], Hall [56], and Weston *et al.* [57]. Similarly, we conclude that Eq. (36) is even stronger than those relations enforced by replacing their lower bound C_{AB} with D_{AB} .

Concluding remarks.—As shown in Supplemental Material the canonical purification method can be used to strengthen the Robertson relation $\sigma(A)\sigma(B) \geq |C_{AB}|$ [58] to obtain

$$\sigma(A)\sigma(B) \geq D_{AB}. \quad (37)$$

Hayashi [59, p. 194] derived this relation by a different method and suggested that C_{AB} in Eq. (4) can be replaced by D_{AB} , but it is not clear whether his method can be used to derive the results obtained in this paper.

The rms error $\varepsilon(A)$ is uniquely derived from the classical notion of root-mean-square error if $U^\dagger(I \otimes M)U$ and $A \otimes I$ commute as in the case of linear position measurements [21]. It is also pointed out in Ref. [47] that $\varepsilon(A)$ coincides with the root-mean-square error of quantum estimation problems [60] for orthogonal families of pure states with the uniform prior distribution, commonly arising in quantum cryptographic protocols. We adopt the state-dependent approach instead of the state-independent approach recently advocated by Busch *et al.* [61], since the properties of the “mean error” can be more suitably described in the state-dependent approach, whereas the state-independent approach is more suitable for describing the “worst case error” [62]. In the state-dependent approach, the precise measurement of an observable A in a given state ρ is characterized in terms of the rms error as the one which satisfies $\varepsilon(A) = 0$ for all ϕ in the cyclic subspace spanned by A and ρ [63–65]. A modification $\bar{\varepsilon}(A)$ of $\varepsilon(A)$ was proposed in [65] to satisfy the condition that $\bar{\varepsilon}(A) = 0$ if and only if the observable A is precisely measured in the state ρ , to clear a problem raised by Busch *et al.* [66]. Then, $\bar{\varepsilon}(A)$ satisfies all the relations obtained in this paper, by the relation $\bar{\varepsilon}(A) \geq \varepsilon(A)$. The definition of $\eta(B)$ is derived analogously, although there are continuing debates on alternative approaches [21, 57, 61, 62]. Further discussions on the significance of the state-dependent approach would be out of the scope of this paper and will be given elsewhere.

There has been a controversy [67, 68] on the question about experimental accessibility of the error $\varepsilon(A)$ and disturbance $\eta(B)$. To clear this question two methods have been proposed so far: the “three-state method” proposed in Ref. [10] and the “weak-measurement method” proposed by Lund-Wiseman [49] based on the relation between the rms error/disturbance and the weak joint probability given in Refs. [18, 30, 63]. Those methods have been experimentally demonstrated in Refs. [50–55]. The third method using “two-point quantum correlator” has been proposed recently [69]. We can expect that those methods will observe the new relations obtained in this paper holding even in mixed states as well as their tightness for spin measurements.

The author thanks Francesco Buscemi for helpful discussions. This work was supported by the MIC SCOPE, No. 121806010, by the John Templeton Foundations, No. 35771, and by the JSPS KAKENHI, No. 26247016.

-
- [1] W. Heisenberg, Z. Phys. **43**, 172 (1927).
 - [2] E. B. Davies and J. T. Lewis, Commun. Math. Phys. **17**, 239 (1970).
 - [3] E. B. Davies, *Quantum Theory of Open Systems* (Academic, London, 1976).
 - [4] M. Ozawa, J. Math. Phys. **25**, 79 (1984).
 - [5] M. Ozawa, J. Math. Phys. **27**, 759 (1986).
 - [6] M. Ozawa, in *Squeezed and Nonclassical Light*, edited by P. Tombesi and E. R. Pike (Plenum, New York, 1989), pp. 263–286.
 - [7] P. Busch, P. J. Lahti, and P. Mittelstaedt, *The Quantum Theory of Measurement*, Lecture Notes in Physics m2 (Springer, Berlin, 1991).
 - [8] V. B. Braginsky and F. Y. Khalili, *Quantum Measurement* (Cambridge University Press, Cambridge, 1992).
 - [9] M. Ozawa, Phys. Rev. A **62**, 062101 (2000).
 - [10] M. Ozawa, Ann. Phys. (N.Y.) **311**, 350 (2004).
 - [11] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, NJ, 1955), [English translation of *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin, 1932)].
 - [12] D. Bohm, *Quantum Theory* (Prentice-Hall, New York, 1951).
 - [13] A. Messiah, *Mécanique Quantique*, vol. I (Dunod, Paris, 1959), [*Quantum Mechanics*, Vol. I (North-Holland, Amsterdam, 1959)].
 - [14] L. I. Schiff, *Quantum Mechanics* (MacGraw-Hill, New York, 1968).
 - [15] V. B. Braginsky, Y. I. Vorontsov, and K. S. Thorne, Science **209**, 547 (1980).
 - [16] H. P. Yuen, Phys. Rev. Lett. **51**, 719 (1983), [see also *ibid.* p. 1603].
 - [17] C. M. Caves, Phys. Rev. Lett. **54**, 2465 (1985).
 - [18] M. Ozawa, Phys. Rev. Lett. **60**, 385 (1988).
 - [19] J. Maddox, Nature **331**, 559 (1988).
 - [20] M. Ozawa, Phys. Lett. A **299**, 1 (2002).
 - [21] M. Ozawa, *Disproving Heisenberg's error-disturbance relation* (2013), arXiv:1308.3540v1 [quant-ph].
 - [22] V. Giovannetti, S. Lloyd, and L. Maccone, Science **306**, 1330 (2004).
 - [23] M. Ozawa, Phys. Lett. A **318**, 21 (2003).
 - [24] M. Ozawa, Phys. Rev. A **67**, 042105 (2003).
 - [25] M. Ozawa, Int. J. Quant. Inf. **1**, 569 (2003).
 - [26] M. Ozawa, Phys. Lett. A **320**, 367 (2004).
 - [27] M. Ozawa, J. Opt. B: Quantum Semiclass. Opt. **7**, S672 (2005).
 - [28] E. Arthurs and M. S. Goodman, Phys. Rev. Lett. **60**, 2447 (1988).
 - [29] M. G. Raymer, Am. J. Phys. **62**, 986 (1994).
 - [30] M. Ozawa, in *Quantum Aspects of Optical Communications*, edited by C. Bendjaballah, O. Hirota, and S. Reynaud (Springer, Berlin, 1991), pp. 3–17.
 - [31] S. Ishikawa, Rep. Math. Phys. **29**, 257 (1991).
 - [32] M. Ozawa, AIP Conf. Proc. **734**, 95 (2004).
 - [33] M. Ozawa, Phys. Rev. Lett. **88**, 050402 (2002).
 - [34] P. Busch and L. Loveridge, Phys. Rev. Lett. **106**, 110406 (2011).
 - [35] E. P. Wigner, Z. Phys. **133**, 101 (1952).
 - [36] H. Araki and M. M. Yanase, Phys. Rev. **120**, 622 (1960).
 - [37] M. M. Yanase, Phys. Rev. **123**, 666 (1961).
 - [38] M. Ozawa, Phys. Rev. Lett. **67**, 1956 (1991).
 - [39] M. Ozawa, Phys. Rev. Lett. **89**, 057902 (2002).
 - [40] J. Gea-Banacloche, Phys. Rev. Lett. **89**, 022308 (2002).
 - [41] M. Ozawa, in *Proceedings of the Sixth International Conference on Quantum Communication, Measurement and Computing*, edited by J. H. Shapiro and O. Hirota (Rinton Press, Princeton, 2003), pp. 175–180.
 - [42] J. Gea-Banacloche and M. Ozawa, J. Opt. B: Quantum Semiclass. Opt. **7**, S326 (2005).
 - [43] J. Gea-Banacloche and M. Ozawa, Phys. Rev. A **74**, 060301(R) (2006).
 - [44] T. Karasawa and M. Ozawa, Phys. Rev. A **75**, 032324 (2007).
 - [45] T. Karasawa, J. Gea-Banacloche, and M. Ozawa, J. Phys. A: Math. Theor. **42**, 225303 (2009).
 - [46] C. Branciard, Proc. Natl. Acad. Sci. USA **110**, 6742 (2013).
 - [47] F. Buscemi, M. J. Hall, M. Ozawa, and M. M. Wilde, Phys. Rev. Lett. **112**, 050401 (2014).
 - [48] C. Branciard, Phys. Rev. A **89**, 022124 (2014).
 - [49] A. P. Lund and H. M. Wiseman, New J. Phys. **12**, 093011 (2010).
 - [50] J. Erhart, S. Sponar, G. Sulyok, G. Badurek, M. Ozawa, and Y. Hasegawa, Nature Phys. **8**, 185 (2012).
 - [51] L. A. Rozema, A. Darabi, D. H. Mahler, A. Hayat, Y. Soudagar, and A. M. Steinberg, Phys. Rev. Lett. **109**, 100404 (2012).
 - [52] S.-Y. Baek, F. Kaneda, M. Ozawa, and K. Edamatsu, Sci. Rep. **3**, 2221 (2013).
 - [53] G. Sulyok, S. Sponar, J. Erhart, G. Badurek, M. Ozawa, and Y. Hasegawa, Phys. Rev. A **88**, 022110 (2013).
 - [54] F. Kaneda, S.-Y. Baek, M. Ozawa, and K. Edamatsu, Phys. Rev. Lett. **112**, 020402 (2014).
 - [55] M. Ringbauer, D. N. Biggerstaff, M. A. Broome, A. Fedrizzi, C. Branciard, and A. G. White, Phys. Rev. Lett. **112**, 020401 (2014).
 - [56] M. J. W. Hall, Phys. Rev. A **69**, 052113 (2004).
 - [57] M. M. Weston, M. J. W. Hall, M. S. Palsson, H. M. Wiseman, and G. J. Pryde, Phys. Rev. Lett. **110**, 220402 (2013).
 - [58] H. P. Robertson, Phys. Rev. **34**, 163 (1929).
 - [59] M. Hayashi, *Quantum Information: An Introduction* (Springer, Berlin, 2006).
 - [60] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).
 - [61] P. Busch, P. Lahti, and R. F. Werner, Phys. Rev. Lett. **111**, 160405 (2013).
 - [62] L. A. Rozema, D. H. Mahler, A. Hayat, and A. M. Steinberg, *A note on different definitions of momentum disturbance* (2013), arXiv:1307.3604v1 [quant-ph].
 - [63] M. Ozawa, Phys. Lett. A **335**, 11 (2005).
 - [64] M. Ozawa, Ann. Phys. (N.Y.) **321**, 744 (2006).
 - [65] M. Ozawa, Proc. SPIE **6244**, 62440Q (2006).
 - [66] P. Busch, T. Heinonen, and P. Lahti, Phys. Lett. A **320**, 261 (2004).
 - [67] R. F. Werner, Quantum Inf. Comput. **4**, 546 (2004).
 - [68] K. Koshino and A. Shimizu, Phys. Rep. **412**, 191 (2005).
 - [69] F. Buscemi, M. Dall'Arno, M. Ozawa, and V. Vedral, *Direct observation of any two-point quantum correlation function* (2013), arXiv:1312.4240 [quant-ph].
 - [70] C. F. Gauss, *Theoria Combinationis Observationum Erroribus Minimis Obnoxiae, Pars Prior, Pars Posterior, Supplementum (Societati Regiae Scientiarum Exhibita, Feb. 15, 1821) [Theory of the Combination of Observations Least Subject to Errors: Part One, Part Two, Supplement, translated by G. W. Stewart]*

(Society for Industrial and Applied Mathematics, Philadelphia, USA, 1995).

- [71] E. H. Kennard, Z. Phys. **44**, 326 (1927).
- [72] The commutator $[Q, P]$ is defined by $[Q, P] = QP - PQ$.
- [73] Here, “mean error” is naturally understood as a notion following that of “root-mean-square error” introduced by Gauss [70], who introduced and called it as the “mean error” or the “mean error to be feared” in 1821.
- [74] The standard deviation is defined for any observable A by $\sigma(A)^2 = \langle A^2 \rangle - \langle A \rangle^2$, where $\langle \dots \rangle$ stands for the mean value in a given state.
- [75] In Ref. [1, p. 69] Heisenberg actually derived the relation $\tilde{\sigma}(Q)\tilde{\sigma}(P) = \hbar$ for $\tilde{\sigma}(Q) = \sqrt{2}\sigma(Q)$ and $\tilde{\sigma}(P) = \sqrt{2}\sigma(P)$ in Gaussian wave functions. Kennard [71] derived the relation $\tilde{\sigma}(Q)\tilde{\sigma}(P) \geq \hbar$ that generalizes Heisenberg’s relation to arbitrary wave functions. In the text, we elaborate Heisenberg’s original argument with Kennard’s general relation.
- [76] It follows immediately from (AR) and Eq. (3) that every simultaneous measurement of Q and P with error limit $(\varepsilon(Q), \varepsilon(P))$ satisfies the relation $\varepsilon(Q)\varepsilon(P) \geq \frac{\hbar}{2}$. Then, Eq. (2) follows immediately from this relation, since if Q can be measured with error-disturbance limit $(\varepsilon(Q), \eta(P))$, then Q and P can be simultaneously measured with error limit $(\varepsilon(Q), \eta(P))$ (see e.g., [24]).
- [77] Von Neumann [11, pp. 238–239] wrote “We are then to show that if P, Q are two canonically conjugate quantities, and a system is in a state in which the value of P can be given with the accuracy $\varepsilon\{= \sigma(P)\}$ (i.e., by a P measurement with an error range $\varepsilon\{= \varepsilon(P)\}$), then Q can be known with no greater accuracy than $\eta\{= \sigma(Q)\} = \hbar/[4\pi\varepsilon]$. Or: a measurement of P with the accuracy $\varepsilon\{= \varepsilon(P)\}$ must bring about an indeterminacy $\eta\{= \eta(Q)\} = \hbar/[4\pi\varepsilon]$ in the value of Q .” Terms in $\{\dots\}$ are supplemented by the present author.
- [78] The repeatability hypothesis is formulated as: “If the physical quantity A is measured twice in succession in a system S , then we get the same value each time (Ref. [11], pp. 335)” Under this hypothesis, any precise measurement of A changes the state to be an eigenstate of the measured observable A , which satisfies $\sigma(A) = 0$.
- [79] Davies and Lewis [2, p. 239] wrote “One of the crucial notions is that of repeatability which we show is implicitly assumed in most of the axiomatic treatments of quantum mechanics, but whose abandonment leads to a much more flexible approach to measurement theory.”

Supplemental Material

BRANCIARD'S GEOMETRIC INEQUALITIES

Let \mathcal{E} be a real inner product space. Branciard [46] proved the following three relations hold for any vectors

$\mathbf{a}, \mathbf{b}, \mathbf{m}, \mathbf{n}, \bar{\mathbf{m}}, \bar{\mathbf{n}} \in \mathcal{E}$ with $\mathbf{m} \perp \mathbf{n}$, $\bar{\mathbf{m}} = \mathbf{m}/\|\mathbf{m}\|$, and $\bar{\mathbf{n}} = \mathbf{n}/\|\mathbf{n}\|$.

$$[\|\mathbf{a}\|^2 - (\mathbf{a}, \bar{\mathbf{m}})^2]\|\mathbf{b}\|^2 + \|\mathbf{a}\|^2[\|\mathbf{b}\|^2 - (\mathbf{b}, \bar{\mathbf{n}})^2] + 2\sqrt{\|\mathbf{a}\|^2 - (\mathbf{a}, \bar{\mathbf{m}})^2}\sqrt{\|\mathbf{b}\|^2 - (\mathbf{b}, \bar{\mathbf{n}})^2}\sqrt{\|\mathbf{a}\|^2\|\mathbf{b}\|^2 - (\mathbf{a}, \mathbf{b})^2} \geq (\mathbf{a}, \mathbf{b})^2. \quad (\text{S1})$$

$$\|\mathbf{a} - \mathbf{m}\|^2\|\mathbf{b}\|^2 + \|\mathbf{a}\|^2\|\mathbf{b} - \mathbf{n}\|^2 + 2\|\mathbf{a} - \mathbf{m}\|\|\mathbf{b} - \mathbf{n}\|\sqrt{\|\mathbf{a}\|^2\|\mathbf{b}\|^2 - (\mathbf{a}, \mathbf{b})^2} \geq (\mathbf{a}, \mathbf{b})^2. \quad (\text{S2})$$

STRENGTHENING OF ERROR-DISTURBANCE RELATIONS FOR MIXED STATES

In what follows we shall complete the proof of Eq. (26).

Let \mathcal{H}^* be the dual space of \mathcal{H} . Then, \mathcal{H}^* consists of all bra vectors $\langle\psi| \in \mathcal{H}^*$ corresponding to all ket vectors $|\psi\rangle \in \mathcal{H}$. The inner product on \mathcal{H}^* is such that

$$(\langle\xi|, \langle\eta|) = \langle\eta|\xi\rangle$$

for all ket vectors $|\xi\rangle, |\eta\rangle \in \mathcal{H}$.

For any operator A on \mathcal{H} , we define the operator A^* on \mathcal{H}^* by

$$A^*\langle\eta| = \langle\eta|A$$

for all $|\eta\rangle \in \mathcal{H}$. Then, we have

$$(\langle\xi|, A^*\langle\eta|) = \langle\eta|A|\xi\rangle$$

for all $|\xi\rangle, |\eta\rangle \in \mathcal{H}$.

Now we suppose $\mathcal{H}' = \mathcal{H}^*$. The “canonical purification” $|\Psi\rangle \in \mathcal{H} \otimes \mathcal{H}^*$ of

$$\rho = \sum_j p_j |\phi_j\rangle\langle\phi_j|$$

is defined by

$$|\Psi\rangle = \sum_j \sqrt{p_j} |\phi_j\rangle \otimes \langle\phi_j|,$$

which satisfies

$$\rho = \text{Tr}_{\mathcal{H}^*} [|\Psi\rangle\langle\Psi|].$$

The constant D_{AB} is defined by

$$D_{AB} = \frac{1}{2} \text{Tr}(|\sqrt{\rho}[A, B]\sqrt{\rho}|).$$

Let

$-i\sqrt{\rho}[A, B]\sqrt{\rho} = W^\dagger |\sqrt{\rho}[A, B]\sqrt{\rho}|$ be a polar decomposition of the operator $-i\sqrt{\rho}[A, B]\sqrt{\rho}$. Since $-i\sqrt{\rho}[A, B]\sqrt{\rho}$ is self-adjoint, W can be chosen as a self-adjoint unitary operator on \mathcal{H} . Then, W^* is a self-adjoint unitary operator on \mathcal{H}^* . We extend the observable B on \mathcal{H} to the observable B'_W on $\mathcal{H} \otimes \mathcal{H}^*$ by

$$B'_W = (B - \beta I_{\mathcal{H}}) \otimes W^*,$$

where $\beta = \text{Tr}[B\rho]$. Then, we have

$$\sigma(B)^2 = \langle\Psi|B'^2_W|\Psi\rangle \geq \sigma(B'_W)^2.$$

As in the main text, A' and U' are given by

$$\begin{aligned} A' &= A \otimes I_{\mathcal{H}'}, \\ U' &= U \otimes I_{\mathcal{H}'}. \end{aligned}$$

Let

$$|\Psi'\rangle = |\Psi\rangle \otimes |\xi\rangle.$$

We also have

$$\begin{aligned}
\eta(B'_W) &= \|B'_W(\Delta t)|\Psi'\rangle - B'_W(0)|\Psi'\rangle\| \\
&= \|U^\dagger[(B - \beta I_{\mathcal{H}}) \otimes W^* \otimes I_{\mathcal{K}}]U'|\Psi'\rangle - (B - \beta I_{\mathcal{H}}) \otimes W^* \otimes I_{\mathcal{K}}|\Psi'\rangle\| \\
&= \|U^\dagger[(B - \beta I_{\mathcal{H}}) \otimes I_{\mathcal{K}}]U \otimes W^*|\Psi'\rangle - (B - \beta I_{\mathcal{H}}) \otimes I_{\mathcal{K}} \otimes W^*|\Psi'\rangle\| \\
&= \|[U^\dagger(B \otimes I_{\mathcal{K}})U - B \otimes I_{\mathcal{K}}] \otimes W^*|\Psi'\rangle\| \\
&= \|I_{\mathcal{H}} \otimes I_{\mathcal{K}} \otimes W^*[U^\dagger(B \otimes I_{\mathcal{K}})U - B \otimes I_{\mathcal{K}}] \otimes I_{\mathcal{H}^*}|\Psi'\rangle\| \\
&= \|[U^\dagger(B \otimes I_{\mathcal{K}})U - B \otimes I_{\mathcal{K}}] \otimes I_{\mathcal{H}^*}|\Psi'\rangle\| \\
&= \text{Tr}\{[U^\dagger(B \otimes I_{\mathcal{K}})U - B \otimes I_{\mathcal{K}}]^2 \rho \otimes |\xi\rangle\langle\xi|\}^{1/2} \\
&= \eta(B).
\end{aligned}$$

For $C_{A'B'_W}$, we have

$$\begin{aligned}
2iC_{A'B'_W} &= \langle\Psi|[A', B'_W]|\Psi\rangle \\
&= \langle\Psi|[A \otimes I_{\mathcal{H}^*}, B \otimes W^*]|\Psi\rangle \\
&= \langle\Psi|[A, B] \otimes W^*|\Psi\rangle \\
&= \left(\sum_j \sqrt{p_j} |\phi_j\rangle \otimes \langle\phi_j|, ([A, B] \otimes W^*) \sum_k \sqrt{p_k} |\phi_k\rangle \otimes \langle\phi_k| \right) \\
&= \sum_{j,k} \sqrt{p_j} \sqrt{p_k} \left(|\phi_j\rangle \otimes \langle\phi_j|, [A, B] |\phi_k\rangle \otimes W^* \langle\phi_k| \right) \\
&= \sum_{j,k} \sqrt{p_j} \sqrt{p_k} \left(|\phi_j\rangle, [A, B] |\phi_k\rangle \right) \left(\langle\phi_j|, W^* \langle\phi_k| \right) \\
&= \sum_{j,k} \sqrt{p_j} \sqrt{p_k} \langle\phi_j|[A, B]|\phi_k\rangle \langle\phi_k|W|\phi_j\rangle \\
&= \sum_j \sqrt{p_j} \langle\phi_j|[A, B] \left(\sum_k \sqrt{p_k} |\phi_k\rangle \langle\phi_k| \right) W |\phi_j\rangle \\
&= \sum_j \sqrt{p_j} \langle\phi_j|[A, B] \sqrt{\rho} W |\phi_j\rangle \\
&= \text{Tr}[W \sqrt{\rho} [A, B] \sqrt{\rho}] \\
&= i\text{Tr}[\sqrt{\rho} [A, B] \sqrt{\rho}] \\
&= 2iD_{AB}.
\end{aligned}$$

Thus, in addition to

$$\begin{aligned}
\sigma(A') &= \sigma(A), \\
\varepsilon(A') &= \varepsilon(A),
\end{aligned}$$

we have

$$\begin{aligned}
\sigma(B'_W) &\leq \sigma(B), \\
\eta(B'_W) &= \eta(B), \\
C_{A'B'_W} &= D_{AB}.
\end{aligned}$$

Since from Eq. (13) we have

$$\begin{aligned}
&\varepsilon(A')^2 \sigma(B'_W)^2 + \sigma(A')^2 \eta(B'_W)^2 \\
&\quad + 2\varepsilon(A') \eta(B'_W) \sqrt{\sigma(A')^2 \sigma(B'_W)^2 - C_{A'B'_W}^2} \geq C_{A'B'_W}^2,
\end{aligned}$$

we conclude Eq. (26):

$$\begin{aligned}
&\varepsilon(A)^2 \sigma(B)^2 + \sigma(A)^2 \eta(B)^2 \\
&\quad + 2\varepsilon(A) \eta(B) \sqrt{\sigma(A)^2 \sigma(B)^2 - D_{AB}^2} \geq D_{AB}^2.
\end{aligned}$$

ERROR-TRADEOFF RELATION

In what follows, we shall drive Eq. (S5), the error-tradeoff version of Eq. (26), from Eq. (S2).

Let \mathcal{H} be a Hilbert space describing a quantum system \mathbf{S} . A *joint measurement model* for \mathcal{H} is a quadruple $(\mathcal{K}, |\xi\rangle, \mathcal{A}, \mathcal{B})$ consisting of a Hilbert space \mathcal{K} , a unit vector $|\xi\rangle \in \mathcal{K}$, and mutually commuting self-adjoint operators \mathcal{A}, \mathcal{B} on $\mathcal{H} \otimes \mathcal{K}$. For any pair of observables A, B and a density operator ρ on \mathcal{H} , the *rms errors* $\varepsilon(A, \mathcal{A}, \rho)$ and $\varepsilon(B, \mathcal{B}, \rho)$ for joint A, B

measurement by $(\mathcal{K}, |\xi\rangle, \mathcal{A}, \mathcal{B})$ in ρ are defined by

$$\varepsilon(A, \mathcal{A}, \rho) = \text{Tr}[(\mathcal{A} - A \otimes I_{\mathcal{K}})^2 \rho \otimes |\xi\rangle\langle\xi|]^{1/2}, \quad (\text{S3})$$

$$\varepsilon(B, \mathcal{B}, \rho) = \text{Tr}[(\mathcal{B} - B \otimes I_{\mathcal{K}})^2 \rho \otimes |\xi\rangle\langle\xi|]^{1/2}. \quad (\text{S4})$$

In what follows we shall prove the following theorem.

Theorem 1. *Let A, B be a pair of observables on \mathcal{H} and let ρ be a density operator on \mathcal{H} . Any joint measurement model $(\mathcal{K}, |\xi\rangle, \mathcal{A}, \mathcal{B})$ for \mathcal{H} satisfy the relation*

$$\begin{aligned} & \varepsilon(A)^2 \sigma(B)^2 + \sigma(A)^2 \varepsilon(B)^2 \\ & + 2\varepsilon(A)\varepsilon(B)\sqrt{\sigma(A)^2 \sigma(B)^2 - D_{AB}^2} \geq D_{AB}^2, \end{aligned} \quad (\text{S5})$$

where $\varepsilon(A) = \varepsilon(A, \mathcal{A}, \rho)$ and $\varepsilon(B) = \varepsilon(B, \mathcal{B}, \rho)$, and $\sigma(A) = \sigma(A, \rho)$ and $\sigma(B) = \sigma(B, \rho)$ are the standard deviations of A, B in ρ .

Proof. Let $\mathcal{L}(\mathcal{W})$ be the space of liner operators on a Hilbert space \mathcal{W} ; if \mathcal{W} is infinite dimensional, we should consider the space $\mathcal{L}_{HS}(\mathcal{W})$ of ‘‘Hilbert-Schmidt class’’ operators A satisfying $\text{Tr}[A^*A] < \infty$ instead of $\mathcal{L}(\mathcal{W})$, but for simplicity we abuse the notation $\mathcal{L}(\mathcal{W})$ to denote $\mathcal{L}_{HS}(\mathcal{W})$ even in that case.

For any $X, Y \in \mathcal{L}(\mathcal{W})$, define the real number (X, Y) by

$$(X, Y) = \text{ReTr}(X^\dagger Y). \quad (\text{S6})$$

Then, we have the following:

(i) (Linearity) $(Z, xX + yY) = \text{ReTr}Z^\dagger(xX + yY) = x\text{ReTr}Z^\dagger X + y\text{ReTr}Z^\dagger Y = x(Z, X) + y(Z, Y)$.

(ii) (Symmetry) $(Y, X) = \text{ReTr}Y^\dagger X = \text{Re}(\text{Tr}Y^\dagger X)^* = \text{ReTr}X^\dagger Y = (X, Y)$.

(iii) (Positivity) $(X, X) = \text{ReTr}[X^\dagger X] \geq 0$

(iv) (Non-degeneracy) If $(X, X) = 0$, then $\text{Tr}[X^\dagger X] = 0$, so that $X = 0$.

Thus, $\mathcal{L}(\mathcal{W})$ is a real linear space with real-valued inner product (X, Y) and norm $\|X\| = (X, X)^{1/2}$ for $X, Y \in \mathcal{L}(\mathcal{W})$.

Let A, B be a pair of observables on \mathcal{H} and let ρ be a density operator on \mathcal{H} . Let $(\mathcal{K}, |\xi\rangle, \mathcal{A}, \mathcal{B})$ be a joint measurement model for \mathcal{H} . Let W be a self-adjoint unitary operator in \mathcal{H} satisfying the polar decomposition formula

$$-iW\sqrt{\rho}[A, B]\sqrt{\rho} = |\sqrt{\rho}[A, B]\sqrt{\rho}|.$$

Now, we define vectors $\mathbf{a}, \mathbf{b}, \mathbf{m}, \mathbf{n} \in \mathcal{L}(\mathcal{W})$ for $\mathcal{W} = \mathcal{H} \otimes \mathcal{K}$ by

$$\mathbf{a} = A_0\sqrt{\rho} \otimes |\xi\rangle\langle\xi|, \quad (\text{S7})$$

$$\mathbf{b} = -iB_0\sqrt{\rho}W \otimes |\xi\rangle\langle\xi|, \quad (\text{S8})$$

$$\mathbf{m} = \mathcal{A}_0(\sqrt{\rho} \otimes |\xi\rangle\langle\xi|), \quad (\text{S9})$$

$$\mathbf{n} = -i\mathcal{B}_0(\sqrt{\rho}W \otimes |\xi\rangle\langle\xi|), \quad (\text{S10})$$

where

$$A_0 = A - \text{Tr}[A\rho]I_{\mathcal{H}}, \quad (\text{S11})$$

$$B_0 = B - \text{Tr}[B\rho]I_{\mathcal{H}}, \quad (\text{S12})$$

$$\mathcal{A}_0 = A - \text{Tr}[A\rho]I_{\mathcal{H} \otimes \mathcal{K}}, \quad (\text{S13})$$

$$\mathcal{B}_0 = B - \text{Tr}[B\rho]I_{\mathcal{H} \otimes \mathcal{K}}. \quad (\text{S14})$$

Then, we have

$$\begin{aligned} (\mathbf{a}, \mathbf{b}) &= \text{ReTr}[(A_0\sqrt{\rho})^\dagger(-i)B_0\sqrt{\rho}W] \\ &= \text{Re}(-i\text{Tr}[W\sqrt{\rho}A_0B_0\sqrt{\rho}]) \\ &= \text{ImTr}[W\sqrt{\rho}A_0B_0\sqrt{\rho}] \\ &= \frac{1}{2i}\text{Tr}(W\sqrt{\rho}[A_0, B_0]\sqrt{\rho}) \\ &= \frac{1}{2i}\text{Tr}(W\sqrt{\rho}[A, B]\sqrt{\rho}) \\ &= \frac{1}{2}\text{Tr}(|\sqrt{\rho}[A, B]\sqrt{\rho}|). \end{aligned}$$

Thus, we obtain

$$(\mathbf{a}, \mathbf{b}) = D_{AB}. \quad (\text{S15})$$

From

$$\begin{aligned} (\mathbf{m}, \mathbf{n}) &= \text{Re}\{-i\text{Tr}[\mathcal{A}_0\mathcal{B}_0(\sqrt{\rho}W\sqrt{\rho} \otimes |\xi\rangle\langle\xi|)]\} \\ &= \text{ImTr}[\mathcal{A}_0\mathcal{B}_0(\sqrt{\rho}W\sqrt{\rho} \otimes |\xi\rangle\langle\xi|)] \\ &= \frac{1}{2i}\text{Tr}\{\mathcal{A}_0, \mathcal{B}_0\}(\sqrt{\rho}W\sqrt{\rho} \otimes |\xi\rangle\langle\xi|) \\ &= 0 \end{aligned}$$

we have

$$\mathbf{m} \perp \mathbf{n}. \quad (\text{S16})$$

We also have the following relations.

$$\|\mathbf{a}\|^2 = \text{Tr}[|A_0\sqrt{\rho}|^2] = \sigma(A, \rho)^2, \quad (\text{S17})$$

$$\|\mathbf{b}\|^2 = \text{Tr}[|iB_0\sqrt{\rho}W|^2] = \sigma(B, \rho)^2, \quad (\text{S18})$$

$$\begin{aligned} \|\mathbf{m} - \mathbf{a}\|^2 &= \text{Tr}[(\mathcal{A}_0 - A_0 \otimes I_{\mathcal{K}})\sqrt{\rho} \otimes |\xi\rangle\langle\xi|]^2 \\ &= \varepsilon(A, \mathcal{A}, \rho)^2, \end{aligned} \quad (\text{S19})$$

$$\begin{aligned} \|\mathbf{n} - \mathbf{b}\|^2 &= \text{Tr}[(\mathcal{B}_0 - B_0 \otimes I_{\mathcal{K}})\sqrt{\rho} \otimes |\xi\rangle\langle\xi|]^2 \\ &= \varepsilon(B, \mathcal{B}, \rho)^2. \end{aligned} \quad (\text{S20})$$

Therefore, we obtain Eq. (S5) from Eq. (S2). \square

Let $(\mathcal{K}, |\xi\rangle, U, M)$ be a measuring process for \mathcal{H} with rms error $\varepsilon(A)$ and rms disturbance $\eta(B)$ in a state ρ . Then, $(\mathcal{K}, |\xi\rangle, M(\Delta t), B(\Delta t))$ is a joint measurement model such that

$$\varepsilon(A, M(\Delta t), \rho) = \varepsilon(A), \quad (\text{S21})$$

$$\varepsilon(B, B(\Delta t), \rho) = \eta(B). \quad (\text{S22})$$

Thus, Eq. (26) can also be derived from Eq. (S5).

Note that from the Schwarz inequality

$$\|\mathbf{a}\| \|\mathbf{b}\| \geq |(\mathbf{a}, \mathbf{b})|$$

with Eqs. (S15), (S17), and (S18) we obtain the relation

$$\sigma(A, \rho)\sigma(B, \rho) \geq D_{AB} \quad (\text{S23})$$

for any observable A, B and state ρ . This strengthens the Robertson inequality [58]

$$\sigma(A, \rho)\sigma(B, \rho) \geq |C_{AB}| \quad (\text{S24})$$

if ρ is a mixed state.

Note also that the tensor product space $\mathcal{W} \otimes \mathcal{W}^*$ is isomorphic to the space $\mathcal{L}(\mathcal{W})$ by the correspondence between $|\xi\rangle \otimes \langle\eta| \in \mathcal{W} \otimes \mathcal{W}^*$ and the operator $T_{|\xi\rangle \otimes \langle\eta|} = |\xi\rangle\langle\eta| \in \mathcal{L}(\mathcal{W})$ such that $T_{|\xi\rangle \otimes \langle\eta|}|\psi\rangle = \langle\eta|\psi\rangle|\xi\rangle$ for all $|\psi\rangle \in \mathcal{W}$. Moreover, for any density operator ρ on \mathcal{W} , its canonical purification $|\Psi\rangle \in \mathcal{W} \otimes \mathcal{W}^*$ corresponds to $\sqrt{\rho} \in \mathcal{L}(\mathcal{W})$ under the above isomorphism. Thus, for any state ρ the proof using the operator $\sqrt{\rho} \in \mathcal{L}(\mathcal{W})$ is essentially transferrable to a proof using the canonical purification $|\Psi\rangle \in \mathcal{W} \otimes \mathcal{W}^*$ of ρ . We prefer the operator representation $\sqrt{\rho} \in \mathcal{L}(\mathcal{W})$ for the canonical purification of ρ because of its mathematical tractability.

ERROR-DISTURBANCE RELATION FOR BINARY MEASUREMENTS

In this section, we shall derive Eq. (27) from Eq. (S1).

First, we consider binary joint measurements with the same spectrum condition.

Theorem 2. *Let A, B be a pair of observables on \mathcal{H} , ρ a density operator on \mathcal{H} , and $(\mathcal{K}, |\xi\rangle, \mathcal{A}, \mathcal{B})$ a joint measurement model for \mathcal{H} . Suppose that $A^2 = B^2 = I_{\mathcal{H}}$, $\mathcal{A}^2 = \mathcal{B}^2 = I_{\mathcal{H} \otimes \mathcal{K}}$, and that $\text{Tr}[A\rho] = \text{Tr}[B\rho] = 0$. Then, we have the relation*

$$\hat{\varepsilon}(A)^2 + \hat{\varepsilon}(B)^2 + 2\hat{\varepsilon}(A)\hat{\varepsilon}(B)\sqrt{1 - D_{AB}^2} \geq D_{AB}^2, \quad (\text{S25})$$

where

$$\begin{aligned} \hat{\varepsilon}(A) &= \varepsilon(A, \mathcal{A}, \rho) \sqrt{1 - \frac{\varepsilon(A, \mathcal{A}, \rho)^2}{4}}, \\ \hat{\varepsilon}(B) &= \varepsilon(B, \mathcal{B}, \rho) \sqrt{1 - \frac{\varepsilon(B, \mathcal{B}, \rho)^2}{4}}. \end{aligned}$$

Proof. Under the assumptions $A^2 = B^2 = I_{\mathcal{H}}$, $\mathcal{A}^2 = \mathcal{B}^2 = I_{\mathcal{H} \otimes \mathcal{K}}$, and $\text{Tr}[A\rho] = \text{Tr}[B\rho] = 0$, the vectors $\mathbf{a}, \mathbf{b}, \mathbf{m}, \mathbf{n}$ in the proof of Theorem 1 satisfy the relations

$$(\mathbf{a}, \mathbf{b}) = D_{AB}, \quad (\text{S26})$$

$$\mathbf{m} \perp \mathbf{n}, \quad (\text{S27})$$

$$\|\mathbf{a}\|^2 = \sigma(A, \rho)^2 = 1, \quad (\text{S28})$$

$$\|\mathbf{b}\|^2 = \sigma(B, \rho)^2 = 1, \quad (\text{S29})$$

$$\|\mathbf{m}\|^2 = \sigma(\mathcal{A}, \rho \otimes |\xi\rangle\langle\xi|)^2 = 1, \quad (\text{S30})$$

$$\|\mathbf{n}\|^2 = \sigma(\mathcal{B}, \rho \otimes |\xi\rangle\langle\xi|)^2 = 1, \quad (\text{S31})$$

$$\begin{aligned} \|\mathbf{m} - \mathbf{a}\|^2 &= \text{Tr}[(\mathcal{A}_0 - A_0)\sqrt{\rho} \otimes |\xi\rangle\langle\xi|]^2 \\ &= \varepsilon(A, \mathcal{A}, \rho)^2, \end{aligned} \quad (\text{S32})$$

$$\begin{aligned} \|\mathbf{n} - \mathbf{b}\|^2 &= \text{Tr}[(\mathcal{B}_0 - B_0)\sqrt{\rho} \otimes |\xi\rangle\langle\xi|]^2 \\ &= \varepsilon(B, \mathcal{B}, \rho)^2. \end{aligned} \quad (\text{S33})$$

By the relation $\|\mathbf{a} - \mathbf{m}\|^2 = 2 - 2(\mathbf{a}, \mathbf{m})$ we have

$$\begin{aligned} \|\mathbf{a}\|^2 - (\mathbf{a}, \mathbf{m})^2 &= \|\mathbf{a} - \mathbf{m}\|^2 - \frac{\|\mathbf{a} - \mathbf{m}\|^4}{4} \\ &= \|\mathbf{a} - \mathbf{m}\|^2 \left(1 - \frac{\|\mathbf{a} - \mathbf{m}\|^2}{4}\right) \\ &= \varepsilon(A, \mathcal{A}, \rho)^2 \left(1 - \frac{\varepsilon(A, \mathcal{A}, \rho)^2}{4}\right). \end{aligned} \quad (\text{S34})$$

Similary, we have

$$\|\mathbf{b}\|^2 - (\mathbf{b}, \mathbf{n})^2 = \varepsilon(B, \mathcal{B}, \rho)^2 \left(1 - \frac{\varepsilon(B, \mathcal{B}, \rho)^2}{4}\right). \quad (\text{S35})$$

Thus, Eq. (S25) follows from Eq. (S1). \square

For the error-disturbance scenario in the text, we conclude Eq. (27). The proof runs as follows. Let A, B be a pair of observables on \mathcal{H} , ρ a density operator on \mathcal{H} , and $(\mathcal{K}, |\xi\rangle, U, M)$ a measuring process for \mathcal{H} with $\varepsilon(A)$ and $\eta(B)$ in ρ . Then, we have a joint measurement model $(\mathcal{K}, |\xi\rangle, M(\Delta t), B(\Delta t))$ such that $\varepsilon(A, M(\Delta t), \rho) = \varepsilon(A)$ and $\varepsilon(B, B(\Delta t), \rho) = \eta(B)$. Since $M^2 = I_{\mathcal{K}}$ and $A^2 = I_{\mathcal{H}}$, we have $M(\Delta t)^2 = B(\Delta t)^2 = I_{\mathcal{H} \otimes \mathcal{K}}$. Thus, Eq. (27) follows from Eq. (S25).

ERROR-TRADEOFF RELATION WITH PARAMETERS FOR BIAS AND OUTPUT FLUCTUATION

In what follows, we shall derive Eq. (36) from Eq. (S1).

Let A, B be a pair of observables on \mathcal{H} and let ρ be a density operator on \mathcal{H} . Let $(\mathcal{K}, |\xi\rangle, \mathcal{A}, \mathcal{B})$ be a joint measurement model for \mathcal{H} . Let W be a self-adjoint unitary operator in \mathcal{H} satisfying the polar decomposition formula

$$-iW\sqrt{\rho}[A, B]\sqrt{\rho} = |\sqrt{\rho}[A, B]\sqrt{\rho}|.$$

Now, we define vectors $\mathbf{a}, \mathbf{b}, \mathbf{m}, \mathbf{n} \in \mathcal{L}(\mathcal{W})$ for $\mathcal{W} = \mathcal{H} \otimes \mathcal{K}$ by

$$\mathbf{a} = (A_0\sqrt{\rho}) \otimes |\xi\rangle\langle\xi|, \quad (\text{S36})$$

$$\mathbf{b} = (-iB_0\sqrt{\rho}W) \otimes |\xi\rangle\langle\xi|, \quad (\text{S37})$$

$$\mathbf{m} = \mathcal{A}_0(\sqrt{\rho} \otimes |\xi\rangle\langle\xi|), \quad (\text{S38})$$

$$\mathbf{n} = -i\mathcal{B}_0(\sqrt{\rho}W \otimes |\xi\rangle\langle\xi|), \quad (\text{S39})$$

where

$$A_0 = A - \text{Tr}[A\rho]I_{\mathcal{H}}, \quad (\text{S40})$$

$$B_0 = B - \text{Tr}[B\rho]I_{\mathcal{H}}, \quad (\text{S41})$$

$$\mathcal{A}_0 = \mathcal{A} - \text{Tr}[\mathcal{A}(\rho \otimes |\xi\rangle\langle\xi|)]I_{\mathcal{H} \otimes \mathcal{K}}, \quad (\text{S42})$$

$$\mathcal{B}_0 = \mathcal{B} - \text{Tr}[\mathcal{B}(\rho \otimes |\xi\rangle\langle\xi|)]I_{\mathcal{H} \otimes \mathcal{K}}. \quad (\text{S43})$$

Then, by similar calculations to the derivations of Eq. (S15) and Eq. (S16), we obtain

$$(\mathbf{a}, \mathbf{b}) = D_{AB}, \quad (\text{S44})$$

$$\mathbf{m} \perp \mathbf{n}. \quad (\text{S45})$$

We also have the following relations.

$$\|\mathbf{a}\|^2 = \text{Tr}(|A_0\sqrt{\rho}|^2) = \sigma(A)^2. \quad (\text{S46})$$

$$\|\mathbf{b}\|^2 = \text{Tr}(|-iB_0\sqrt{\rho}W|^2) = \sigma(B)^2. \quad (\text{S47})$$

$$\|\mathbf{m}\|^2 = \text{Tr}[|A_0(\sqrt{\rho} \otimes |\xi\rangle\langle\xi|)|^2] = \sigma(\mathcal{A})^2. \quad (\text{S48})$$

$$\|\mathbf{n}\|^2 = \text{Tr}[|-iB_0(\sqrt{\rho}W \otimes |\xi\rangle\langle\xi|)|^2] = \sigma(\mathcal{B})^2. \quad (\text{S49})$$

We have

$$\begin{aligned} \|\mathbf{b} - \mathbf{n}\|^2 &= \text{Tr}[|\mathcal{B}_0(\sqrt{\rho} \otimes |\xi\rangle\langle\xi|) - (B_0\sqrt{\rho}) \otimes |\xi\rangle\langle\xi||^2] \\ &= \text{Tr}[|\mathcal{B}(\sqrt{\rho} \otimes |\xi\rangle\langle\xi|) - B\sqrt{\rho} \otimes |\xi\rangle\langle\xi| \\ &\quad - \delta(B)(\sqrt{\rho} \otimes |\xi\rangle\langle\xi|)|^2] \\ &= \varepsilon(B)^2 - \delta(B)^2. \end{aligned} \quad (\text{S50})$$

On the other hand, from

$$\|\mathbf{b} - \mathbf{n}\|^2 = \|\mathbf{b}\|^2 + \|\mathbf{n}\|^2 - 2(\mathbf{b}, \mathbf{n}),$$

we have

$$\begin{aligned} \|\mathbf{b}\|^2 - (\mathbf{b}, \mathbf{n})^2 &= \|\mathbf{b}\|^2 - \left(\frac{\|\mathbf{b}\|^2 + \|\mathbf{n}\|^2 - \|\mathbf{b} - \mathbf{n}\|^2}{2\|\mathbf{n}\|} \right)^2 \\ &= \sigma(B)^2 - \left(\frac{\sigma(B)^2 + \sigma(\mathcal{B})^2 - (\varepsilon(B)^2 - \delta(B)^2)}{2\sigma(\mathcal{B})} \right)^2 \\ &= E_{\sigma(\mathcal{B}), \delta(B)}(B)^2. \end{aligned} \quad (\text{S51})$$

Similarly, we have

$$\|\mathbf{a}\|^2 - (\mathbf{a}, \mathbf{m})^2 = E_{\sigma(\mathcal{A}), \delta(A)}(A)^2. \quad (\text{S52})$$

Therefore, we obtain Eq. (36) from Eq. (S1).